

Liquidity and the multiscaling properties of the volume traded on the stock market

ZOLTÁN EISLER¹(*) and JÁNOS KERTÉSZ^{1,2}

¹ *Department of Theoretical Physics, Budapest University of Technology and Economics
- Budapest, H-1111*

² *Laboratory of Computational Engineering, Helsinki University of Technology - Espoo,
Finland*

PACS. 89.75.-k – Complex systems.

PACS. 89.75.Da – Systems obeying scaling laws.

PACS. 05.40.-a – Fluctuation phenomena, random processes, noise, and Brownian motion.

PACS. 89.65.Gh – Economics; econophysics, financial markets, business and management.

Abstract. – We investigate the correlation properties of transaction data from the New York Stock Exchange. The trading activity $f_i(t)$ of each stock i displays a crossover from weaker to stronger correlations at time scales 60 – 390 minutes. In both regimes, the Hurst exponent H depends logarithmically on the liquidity of the stock, measured by the mean traded value per minute. All multiscaling exponents $\tau(q)$ display a similar liquidity dependence, which clearly indicates the lack of a universal form assumed by other studies. The origin of this behavior is both the long memory in the frequency and the size of consecutive transactions.

Financial markets are self-adaptive complex systems and their understanding requires interdisciplinary research, including the application of concepts and tools of statistical physics. The success of modern statistical physics lies to a large extent in explaining phenomena from phase transitions to far-from-equilibrium processes, where the two key concepts have been scaling and universality. When applied to physical systems, both can rely on a solid foundation: renormalization group theory. But how reliable can insights based on these principles be, if we move on to social or economic systems? According to economists, "physicists [simply] suffer from the belief that there must be universal rules" [1].

The aim of present paper is to point out that the assumption of universality can lead to false conclusions regarding stock market dynamics [2]. We use multifractal analysis – an approach very commonly pursued in econophysics – to point out that the size of the company, or more appropriately the liquidity of its stock, affects the observed characteristics of how it is traded on the market. This dependence is continuous, and therefore it means the absence of universality classes in trading dynamics.

By means of multifractal analysis, we show that: (i) Trading activity records show a crossover from weaker to stronger correlations around the time scale of 1 trading day. (ii)

(*) Email: eisler@maxwell.phy.bme.hu

The strength of correlations above the crossover depends logarithmically on the average trading activity of the stock. (iii) The whole family of $\tau(q)$ multiscaling exponents of trading activity shows a similar variation. (iv) These effects originate from an interplay between the autocorrelations of the frequency and the size of consecutive transactions.

The dataset used in our study was taken from the NYSE TAQ database [3], and it contains the records of all transactions of the New York Stock Exchange in the years 2000 – 2002.

Let us denote the total traded value of a stock i in the time window $[t, t + \Delta t]$ by $f_i^{\Delta t}(t)$. This is calculated as a sum of the values of all transactions for the given stock during $[t, t + \Delta t]$. If $N_i^{\Delta t}(t)$ denotes the number of trades for the stock in the interval, and the value of the n -th trade is $V_i(n)$, then one can write this formally as

$$f_i^{\Delta t}(t) = \sum_{\substack{n \\ t_i(n) \in [t, t + \Delta t]}} V_i(n), \quad (1)$$

where the sum runs for the $N_i^{\Delta t}(t)$ trades in the interval.

Though the returns are known to be only short time correlated, financial data contain different kinds of long-range correlations, examples range from volatility to order flow [4–6]. Records of traded value are no exception from this [7–9], and are most often characterized by the Hurst exponent, or in general, by multifractal spectra [10, 11]. Multifractal models represent a dynamically developing approach in describing financial processes both in conservative finance and econophysics (for a review, see [12]).

Recent studies [7, 13] have shown that the standard deviation, and even higher moments of f exist, thus it is possible, to define the q -th order partition function in the following way:

$$\sigma_i^q(\Delta t) = \left\langle \left| f_i^{\Delta t}(t) - \langle f_i^{\Delta t}(t) \rangle \right|^q \right\rangle \propto \Delta t^{\tau_i(q)}, \quad (2)$$

where $\langle \cdot \rangle$ denotes time averaging. For any fixed stock i , the formula defines a $\tau_i(q)$ set of exponents, indexed by q , and determined by the slopes of eq. (2) on a log-log plot ⁽¹⁾. These are often written in the form $\tau(q) = qH(q)$, and $H = H(2)$ is called the Hurst exponent, while other $H(q)$'s are the generalized Hurst exponents. This family of exponents is closely related to the correlation properties of the data. If $H = 0.5$, the data have no long range correlations, while for $H > 0.5$ ($H < 0.5$) signals have persistent (antipersistent) long range correlations. If $H(q) \equiv H$ is independent of q , the signal is self-affine, while nontrivial q -dependence gives rise to multiscaling or multi-affinity.

Here, we present an analysis of the $\sigma_i^q(\Delta t)$ partition functions. We investigated the 2416 stocks which were continuously listed at NYSE during the years 2000 – 2002, and which had an average turnover $\langle f \rangle$ (mean traded value per minute) of at least 100 USD/min. This ensures that there are no extended periods where the stock is not traded at all, and thus $f(t)$ is well-defined.

For the calculation of $\sigma_i^q(\Delta t)$ we used Detrended Fluctuation Analysis [11]. This method uses piecewise polynomial fits to remove instabilities from the data, and often produces good estimates for $\tau(q)$. We tested the robustness of our estimates to the order of this detrending, and varied the order of the polynomials from 1 to 5, but the results did not change significantly.

Then, we divided the stocks into five groups with respect to $\langle f \rangle$: those with 10^2 USD/min $\leq \langle f \rangle < 10^3$ USD/min, those with 10^3 USD/min $\leq \langle f \rangle < 10^4$ USD/min, ..., and finally 10^6 USD/min $\leq \langle f \rangle$. Then, we averaged the $\sigma_i^q(\Delta t)$ partition functions within each group ⁽²⁾. As an example, the results for $q = 2$ are shown in fig. 1(a).

⁽¹⁾Note that throughout the paper we use 10-base logarithms.

⁽²⁾This averaging procedure decreases the noise present in the data, without affecting our main conclusions.

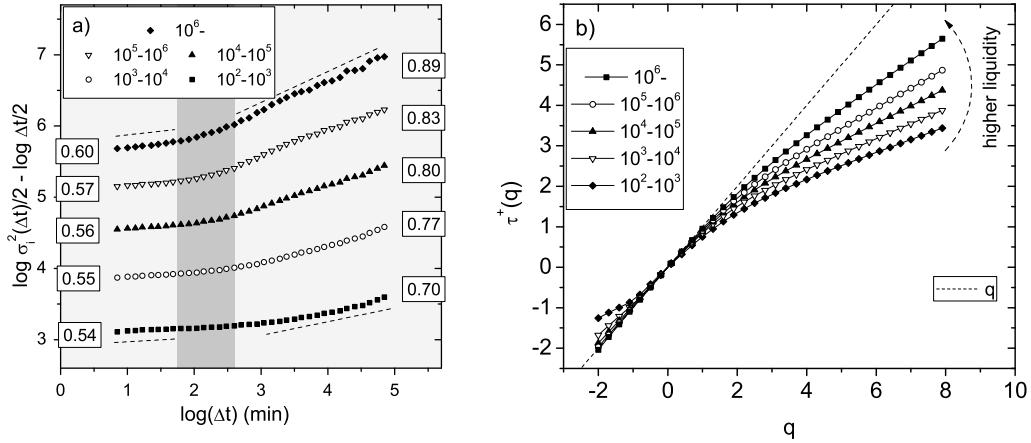


Figure 1 – (a) The normalized partition function $\frac{1}{2} \log \sigma_i^2(\Delta t) - \frac{1}{2} \log \Delta t$ for the five groups of companies. A horizontal line would mean the absence of autocorrelations in the data. Instead, one observes a crossover phenomenon in the regime $\Delta t = 60 - 390$ mins. Below the crossover all groups show weakly correlated behavior. Above the crossover (fluctuations on daily and longer scales), the strength of correlations, and thus the slope corresponding to $H^+ - 0.5$, increases with the liquidity of the stock. Note: Fits are for the regimes below 60 min and above 1700 min. (b) The values of the scaling exponents $\tau^+(q)$, valid for time scales longer than a trading day, for the five groups of companies. Companies with higher average traded value exhibit stronger correlations, and weaker multiscaling than their smaller counterparts. Correspondingly, their $\tau^+(q)$ is greater, and the shape of the curve is closer to the linear relationship $\tau^+(q) = q$.

One finds that, regardless of group, the $\log \sigma^q(\Delta t)$ versus $\log \Delta t$ plots are not straight lines. Instead, one observes a crossover phenomenon ⁽³⁾ [7, 8]: There are two regimes of Δt for which different $\tau(q)$ -s can be identified. For $\Delta t < 60$ min, we are going to use the notation $\tau^-(q)$, while for $\Delta t > 390$ min, $\tau^+(q)$. One can define the related generalized Hurst exponents as $\tau^\pm(q) = qH^\pm(q)$. Systematically, $H^+(q) > H^-(q)$, which means that correlations become stronger when window sizes are greater than 390 min.

Moreover, there is remarkable difference between groups when $\Delta t > 390$ min. This means that the correlations present in the day-to-day variations of trading activity systematically depend on $\langle f \rangle$, as seen from $H^+(q)$ values indicated in fig. 1(a). More of this dependence can be understood if one examines the scaling exponents for more powers of q . This was done by first evaluating the value of $\tau_i^+(q)$ for the independent stocks, and then averaging that for the elements within each group. The results are shown in fig. 1(b). The plot implies, that more liquid stocks (greater $\langle f \rangle$) display stronger correlations than their less liquid (smaller $\langle f \rangle$) counterparts, for any order $q > 0$. This is realized in a way that the degree of multiscaling decreases, and the scaling exponents tend to the fully correlated self-affine behavior with the limiting exponents $\tau^+(q) = q$, $H^+(q) = 1$.

Also note that data were first corrected by the well-known U-shape pattern of daily trading activity (see, e.g., ref. [14]), calculated independently for each group.

⁽³⁾The fact, that the properties of stock market time series are different on time scales shorter than and longer than 1 trading day, was pointed out by many sources. The most common examples are the distribution of returns and the autocorrelations of volatility [4, 5].

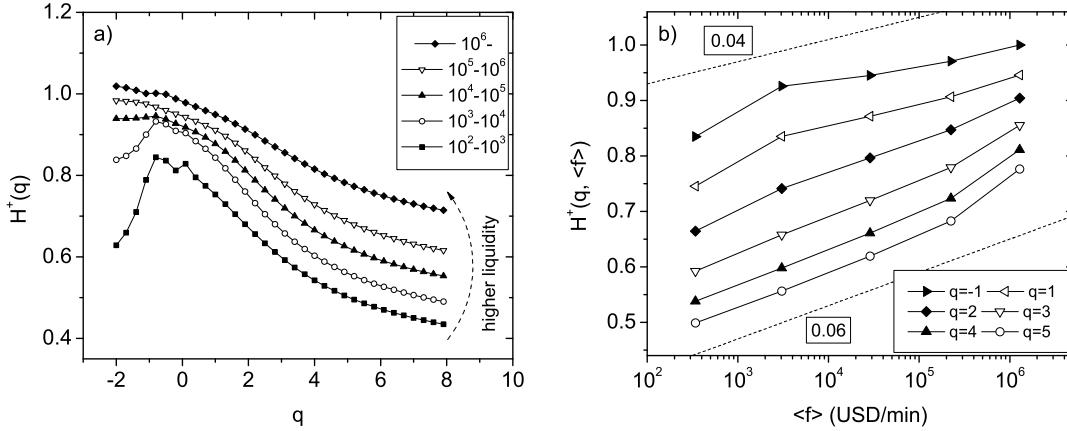


Figure 2 – (a) The values of the generalized Hurst exponents $H^+(q)$, valid for time scales longer than a trading day, for the five groups of companies. The difference in the strength of correlations, and thus $H^+(q)$, is present for all powers q . This implies, that such a dependence on liquidity is present in both low and high trading activity periods. (b) The values of $H^+(q)$, from top to bottom $q = -1, 1, 2, 3, 4, 5$. The points represent the average value for one of the five groups of companies. One can see, that $H^+(q)$ changes in an approximately logarithmic fashion with $\langle f \rangle$. *Note:* Stocks grouped by $\langle f \rangle$, increasing from bottom to top, ranges given in USD/min.

Fig. 2(a) shows the corresponding values of $H^+(q)$. The difference in the $H^+(q)$'s between the groups is present throughout the whole range of q 's, not only for large q 's which are sensitive to the high trading activity. This indicates that the higher level of correlations in more liquid stocks cannot be exclusively attributed to periods of high trading activity. Instead, it is a general phenomenon, that is present continuously⁽⁴⁾.

Despite the presence of non-universality, and that $\tau(q)$ depends on the liquidity of the stock, there is a clear systematic way *how* this dependence is. In fig. 2(b), we plot vertical "cuts" of fig. 2(a). These show, that for a fixed value of q , $\tau^+(q)$ increases with $\langle f \rangle$ in an approximately logarithmic way:

$$\tau^+(q; \langle f \rangle) = C(q) + \gamma(q) \log \langle f \rangle, \quad (3)$$

where $\gamma(q) \approx 0.04 - 0.06$.

Our results imply that the trading of assets of companies with very different size and liquidity cannot be described in a universal manner⁽⁵⁾. There have been studies pointing out such asset-to-asset variations, and the key role of liquidity [16–19], however, they have been consistently overlooked by some econophysics groups. There is a wide range of studies, that calculate ensemble averages over a large number of stocks, irrespective of their liquidities. In some cases universality seems indeed to hold, like for the normalized distribution of returns

⁽⁴⁾One may notice, that there is a strong deviation in the case of stock with low liquidity, and $q < -1$. The origin of this artifact is a finite size effect: The stocks are traded in lots of 100, and thus they cannot be traded in values less than price \times 100. This minimum acts as a cutoff in small fluctuations, to which $q < -1$ moments are very sensitive.

⁽⁵⁾A recent preprint [9] shows similar effects with respect to the market where the stocks are traded. More indications of similar behavior can be found in refs. [7, 15]

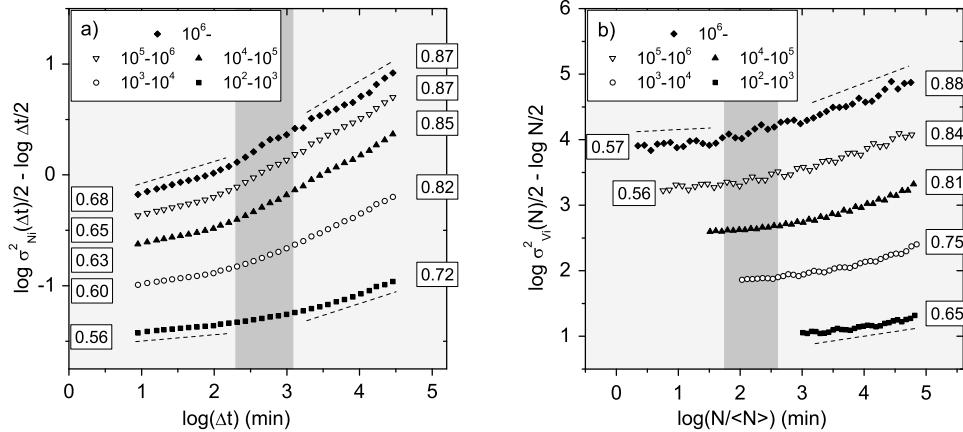


Figure 3 – (a) The normalized partition function $\frac{1}{2} \log \sigma_{N_i}^2(\Delta t) - \frac{1}{2} \log \Delta t$ for the five groups of companies. A horizontal line would mean the absence of autocorrelations in the data. The crossover regime is for slightly longer times, $\Delta t \approx 160 - 1200$ min. Above the crossover the strength of correlations in N , and thus the slopes corresponding to $H_N^+ - 0.5$, increase with the liquidity of the stock. (b) Same as (a), but for $\frac{1}{2} \log \sigma_{V_i}^2(N/\langle N \rangle) - \frac{1}{2} \log N$. The darker shaded area corresponds to the crossover regime of f at $N/\langle N \rangle \approx 60 - 390$ mins. Small stocks are traded infrequently, therefore they have no data points below the crossover. Note: Stocks grouped by $\langle f \rangle$, increasing from bottom to top, ranges given in USD/min. In both plots, fits are for the regimes below 60 min and above 1700 min.

[20, 21]. However, in other cases, as we have just seen, it is misleading to calculate averages for stocks with a wide range of liquidity as done in, e.g., refs. [22–25]. A "typical" $\tau(q)$ or multifractal spectrum of assets is not meaningful in the presence of this clear, systematic dependence.

What aspect of the trading dynamics is the origin of this non-universality? As eq. (1) suggests, the source of fluctuations in f is the fluctuation of N and V (see also ref. [26]). Thus, it is very instructive to define the Hurst exponents of these two processes in analogy with eq. (2). We restrict ourselves to the $q = 2$ moment. One can introduce the H_{N_i} Hurst exponent of the time series $N_i^{\Delta t}(t)$ as

$$\sigma_{N_i}^2(\Delta t) = \left\langle (N_i^{\Delta t}(t) - \langle N_i^{\Delta t} \rangle)^2 \right\rangle \propto \Delta t^{2H_{N_i}}. \quad (4)$$

This H_N describes the temporal correlations of the number of trades.

The results for the group averages of $\sigma_{N_i}^2$, and the asymptotically valid exponents $H_{N_i}^{\pm}$ are shown in fig. 3(a). A comparison with σ_i^2 [fig. 1(a)] shows that both quantities behave similarly: Fluctuations in the number of trades N display crossover and liquidity dependence in the strength of correlations, just like f .

The H_{V_i} Hurst exponent of the so-called tick-by-tick data $V_i(n)$ can be defined as

$$\sigma_{V_i}^2(N/\langle N \rangle) = \left\langle \left(\sum_{n=1}^N V_i(n) - \left\langle \sum_{n=1}^N V_i(n) \right\rangle \right)^2 \right\rangle \propto (N/\langle N \rangle)^{2H_{V_i}}. \quad (5)$$

$\langle f \rangle$ (USD/min)	H_V^-	H^-	H_N^-	H_V^+	H^+	H_N^+
$10^6 -$	0.57	0.60	0.68	0.88	0.89	0.87
$10^5 - 10^6$	0.56	0.57	0.65	0.84	0.83	0.87
$10^4 - 10^5$	–	0.56	0.63	0.81	0.80	0.85
$10^3 - 10^4$	–	0.55	0.60	0.75	0.77	0.82
$10^2 - 10^3$	–	0.54	0.56	0.65	0.70	0.72

Table I – *Hurst exponents for f , N and V for the 5 groups of stocks. For all groups, every exponent is higher above the crossover than below it. Moreover, above the crossover there are no large differences between H_V^+ , H^+ , H_N^+ . From the fits, the errors are estimated to be ± 0.03 . Note: H_V^- is not defined for the 3 groups, whose stocks are not traded at least every 10 minutes.*

The important point here is that the scaling variable is the N number of consecutive trades. This is divided by the $\langle N_i \rangle$ mean number of trades per minute. This is crucial, because the trading frequency of the stocks varies over many orders of magnitudes. Thus N trades corresponds a different time span depending on trading frequency, i.e., on the stock. The scaling variable $N/\langle N_i \rangle$ has a dimension of minutes (just like Δt), and its fixed value always means the same time window size, regardless of the stock.

Moreover, when applying eq. (5), there is a natural lower limit in window size: one cannot take less than one trade, and so $N \geq 1$. Consequently, a group average for $\sigma_{V_i}^2$ is undefined, where the scaling variable would be $N/\langle N_i \rangle < 1/\langle N_i \rangle$ for any stock in the group ⁽⁶⁾. For more liquid stocks, $\langle N_i \rangle$ is larger, thus the minimal window size is smaller.

The results are shown in fig. 3(b). H_V^- is only defined for the two groups, whose stocks are traded at least every 10 minutes, and they indicate weak or no liquidity-dependence. H_V^+ exists for all groups and follows the same trend of increasing correlations for greater liquidity.

The number of transactions in a given time window $[t, t + \Delta t]$ is – to a good approximation – independent from the value of the single transactions ⁽⁷⁾. Under this condition, one can show that for any stock i :

$$\sigma_i^2(\Delta t) = \sigma_{N_i}^2(\Delta t) \langle V_i \rangle^2 + \sigma_{V_i}^2 \left\langle (N_i^{\Delta t})^{2H_{V_i}} \right\rangle, \quad (6)$$

where $\langle V_i \rangle$ is the mean, and $\sigma_{V_i}^2$ is the standard deviation of the value of individual transactions. The origins of the two terms in the formula are the following [26]:

1. The first term describes the effect of fluctuations in the number of transactions. Let us assume, that the size of the transactions is constant, so $V_i(n) = \langle V_i \rangle$, and $\sigma_{V_i}^2 = 0$. Then, the second term is zero, and eq. (6) simplifies to $\sigma_{N_i}^2(\Delta t) \langle V_i \rangle^2$.
2. The second term describes the effect of fluctuations in the value of individual transactions. If one assumes that the number of transactions is the same in every time window, $N_i^{\Delta t}(t) = \langle N_i^{\Delta t} \rangle$, then $\sigma_{N_i}^2 = 0$. The first term becomes zero, and eq. (6) reduces to $\sigma_{V_i}^2 \langle N_i^{\Delta t} \rangle^{2H_{V_i}}$.

Thus the correlations in f originate from the correlations in N and V . By definition, the l.h.s. of eq. (6) is proportional to Δt^{2H_i} . The first term on the r.h.s. is proportional to $\Delta t^{2H_{N_i}}$, while the second term can be estimated to scale as $\Delta t^{2H_{V_i}}$. For large Δt , the behavior of σ^2 is dominated by the larger of H_N and H_V .

⁽⁶⁾We allowed up to 10% of such missing data.

⁽⁷⁾This means, that $N_i^{\Delta t}(t)$ is independent from $f_i^{\Delta t}(t)/N_i^{\Delta t}(t)$. The R^2 values of regressions between the logarithms of these two quantities are typically of the order 0.03 in the data.

This is in agreement with the results summarized in table I. The table also shows that above the crossover there are no major differences between H , H_N and H_V . This means that neither of the two processes dominates in general.

We have studied the correlation functions (2), (4) and (5) found the following: (i) There exists a crossover in the behavior as a function of the time window at $\Delta t \approx 1$ trading day; (ii) There is non-universal (multi)scaling for large Δt with a systematic dependence of the exponents on the liquidity; (iii) Eq. (6) points out the interplay between the fluctuations in the number of trades and the tick-by-tick values, resulting in the observed long term correlations for the activity. While we emphasize the non-universal character of the exponents, we also mean to underline the systematic trends as a function of the company size (liquidity). These properties of trading should be addressed by the future modeling efforts of the stock market.

The authors are grateful to György Andor for his help with financial data. JK is member of the Center for Applied Mathematics and Computational Physics, BME. Support by OTKA T049238 is acknowledged.

References

- [1] BALL P., *Nature*, **441** (2006) 686.
- [2] GALLEGATTI M., KEEN S., LUX T., and ORMEROD P., *to appear in Physica A*, (2006)
- [3] NEW YORK STOCK EXCHANGE, NEW YORK, Trades and Quotes Database for 1993-2003.
- [4] BOUCHAUD J.-P. and POTTERS M., *Theory of Financial Risk*. Cambridge University Press, Cambridge (2000).
- [5] MANTEGNA R. and STANLEY H.E., *Introduction to Econophysics: Correlations and Complexity in Finance* Cambridge University Press (1999).
- [6] ZOVKO I. and FARMER J.D., *Quantitative Finance*, **2** (2002) 387.
- [7] EISLER Z. and KERTÉSZ J., *Eur. Phys. J. B*, **51** (2006) 145 (also physics/0508156)
- [8] IVANOV P., YUEN A., PODOBNIK B., and LEE Y., *Phys. Rev. E*, **69** (2004) 56107.
- [9] YUEN A. and IVANOV P., Impact of stock market microstructure on intertrade time and price dynamics. physics/0508203 (2005).
- [10] VICSEK T., *Fractal Growth Phenomena* World Scientific Publishing (1992).
- [11] KANTELHARDT J., ZSCHIEGNER S., KOSCIELNY-BUNDE E., HAVLIN S., BUNDE A., and STANLEY H.E., *Physica A*, **316** (2002) 87.
- [12] BORLAND L., BOUCHAUD J.-P., MUZY J.-F., and ZUMBACH G., The dynamics of financial markets – Mandelbrot’s multifractal cascades, and beyond. cond-mat/0501292 (2005).
- [13] QUEIRÓS S.M.D., *Europhys. Lett.*, **71** (2005) 339.
- [14] EISLER Z., KERTÉSZ J., YOOK S.-H., and BARABÁSI A.-L., *Europhys. Lett.*, **69** (2005) 664.
- [15] BONANNO G., LILLO F., and MANTEGNA R., *Physica A*, **280** (2000) 136.
- [16] PÁSTOR L. and STAMBAUGH R., *Journal of Political Economy*, **111** (2003) 642.
- [17] CHORDIA T., ROLL R., and SUBRAHMANYAM A., *Journal of Finance*, **56** (2001) 501.
- [18] FARMER J.D., GILLEMET L., LILLO F., MIKE S., and SEN A., *Quantitative Finance*, **4** (2004) 383.
- [19] FARMER J.D. and LILLO F., *Quantitative Finance*, **4** (2004) C7.
- [20] LUX T., *Applied Financial Economics*, **6** (1996) 463.
- [21] GOPIKRISHNAN P., MEYER M., AMARAL L., and STANLEY H.E., *Eur. Phys. J. B*, **3** (1998) 139.
- [22] GOPIKRISHNAN P., PLEROU V., GABAIX X., and STANLEY H.E., *Phys. Rev. E*, **62** (2000) 4493.
- [23] KWAPIEN J., OSWIECIMKA P., and DROZDZ S., *Physica A*, **350** (2005) 466.
- [24] MATIA K., ASHKENAZY Y., and STANLEY H.E., *Europhys. Lett.*, **61** (2003) 422.
- [25] MOYANO L., DE SOUZA J., and QUEIRÓS S.M.D., On the multi-fractal structure of traded volume in financial markets. physics/0512240 (2005)
- [26] EISLER Z. and KERTÉSZ J., *Phys. Rev. E*, **71** (2005) 057104.